# Advances in the simulation and automated measurement of well-sorted granular material: 2. Direct measures of particle properties

D. Buscombe[1] and D. M. Rubin[2]

Received 24 January 2011; revised 26 January 2012; accepted 30 January 2012; published 3 April 2012.

[1] In this, the second of a pair of papers on the structure of well-sorted natural granular material (sediment), new methods are described for automated measurements from images of sediment, of: 1) particle-size standard deviation (arithmetic sorting) with and without apparent void fraction; and 2) mean particle size in material with void fraction. A variety of simulations of granular material are used for testing purposes, in addition to images of natural sediment. Simulations are also used to establish that the effects on automated particle sizing of grains visible through the interstices of the grains at the very surface of a granular material continue to a depth of approximately 4 grain diameters and that this is independent of mean particle size. Ensemble root-mean squared error between observed and estimated arithmetic sorting coefficients for 262 images of natural silts, sands and gravels (drawn from 8 populations) is 31%, which reduces to 27% if adjusted for bias (slope correction between observed and estimated values). These methods allow non-intrusive and fully automated measurements of surfaces of unconsolidated granular material. With no tunable parameters or empirically derived coefficients, they should be broadly universal in appropriate applications. However, empirical corrections may need to be applied for the most accurate results. Finally, analytical formulas are derived for the one-step pore-particle transition probability matrix, estimated from the image's autocorrelogram, from which void fraction of a section of granular material can be estimated directly. This model gives excellent predictions of bulk void fraction yet imperfect predictions of pore-particle transitions.

**Citation:** Buscombe, D., and D. M. Rubin (2012), Advances in the simulation and automated measurement of well-sorted granular material: 2. Direct measures of particle properties, *J. Geophys. Res.*, *117*, F02002, doi:10.1029/2011JF001975.

## 1. Introduction

### 1.1. Objectives

[2] This is the second of a pair of papers (the first being *Buscombe and Rubin* [2012, hereinafter part 1]) which addresses the particle-scale structure of natural granular material. The principal objective here is to provide simple optical techniques for the non-intrusive, calibration-free and automated measurement of some ensemble geometric properties of granular material in a digital image. Part 1 presented an approach for simulating realistic volumes of granular material. We use various outputs from this model, as well as images of real sediment from a variety of sedimentary populations, to develop and test methods for the standard deviation of particle sizes (arithmetic sorting), for sediment both with and without a visible and quantifiable areal (termed 'apparent') void fraction.

[3] In a plan view image of a granular material, as the geometric projection of a 3D surface onto a 2D plane, particles overlap (parts of particles are sitting on top of others, wholly or partially obscuring those underneath). There are no visible voids, only visible surface and parts of subsurface particles (see part 1 for a more detailed discussion). The terminology used in this paper for such surfaces of sediment is a granular material without apparent void fraction. Sections (slices) through a volume of granular material with a bulk porosity will have a real void fraction.

[4] The Fourier-optics principles utilized here follow from *Buscombe et al.* [2010] which provided a similar method for the measurement of mean particle size of touching particles in an image of sediment. Here, that method is extended to images of particles with a significant apparent void fraction (i.e., non-touching particles).

[5] Previous methods have been proposed for estimating the sorting coefficient from the spatial autocorrelation sequence (autocorrelogram) of an image of sediment. *Rubin* [2004] proposed a method based on least squares analysis of an empirically determined autocorrelogram from an image of sediment, and a calibration matrix of autocorrelograms each representing sediment with different and near-uniform

[1]School of Marine Science and Engineering, University of Plymouth, Plymouth, UK.
[2]United States Geological Survey, Santa Cruz, California, USA.

particle size. In order to keep particle-size quantities positive, the least squares problem is solved using a constrained optimization algorithm. *Warrick et al.* [2009] proposed a method based on the method by *Rubin* [2004] and on the observation that the standard error of individual particle sizes computed at each lag was non-linearly related to measured sample geometric sorting coefficients of beach gravels. An empirical power law between these standard errors and sorting was reasonable ($r^2$ = 0.69), but site-specific. It did, however, suggest that the shape of the 1D autocorrelogram calculated from the individual sample could be a potential metric for sample sorting.

[6] However, the present authors found that both methods produce inconsistent results. Although the method of *Rubin* [2004] is theoretically valid, it commonly produces incorrect particle-size distributions, possibly because the least squares problem tends to be ill-posed or possibly because calculated tails of the size distribution are too sensitive to errors in the measured autocorrelation or to the measured size distribution of calibration samples. A calibration-free 'universal' method that works equally well for many different granular materials is preferable.

### 1.2. Paper Structure

[7] The organization of this paper is as follows. First, the various methods are developed in section 2 for estimating sorting and mean particle size directly from an image of granular material (based on the image's autocorrelogram), with no tunable coefficients or other calibration, for granular material with and without an apparent void fraction. This section includes a modification to the *Buscombe et al.* [2010] technique for estimating mean particle size to account for the existence of an apparent void fraction. The materials with an apparent void fraction are essentially 2D sections (slices) made through inherently 3D granular material. The materials without an apparent void fraction are the 'surface images' (which, as explained above, are the projections of a 3D surfaces onto a 2D planes), where the overhead perspective means parts of subsurface particles are viewed through the interstices of the particles at the very surface (however, collectively this image is called a 'surface').

[8] In order to test the techniques contained in section 2, a combination of simulations and real images of sediment was used. Due to the structural complexities of many natural granular materials, it is useful to have realistic simulations with known or well constrained statistical properties. Only simulated material can produce large population sets quickly, and with a range of combinations of particle size, particle shape, packing, porosity and other macroscopic (particle-scale or larger) properties. Section 3 provides a brief summary of how the generalized modeling approach presented in part 1 was used to create sediment surfaces with which to test the algorithms.

[9] For sediment surfaces, measurements of the diameters of completely visible particles, or bulk measurements of the surface layers such as sieving, are not appropriate metrics against which to evaluate the performance of the grain size metrics obtained from methods such as these. *Barnard et al.* [2007] and *Buscombe et al.* [2010] carried out manual point counts of intermediate ('b-axis') particle diameters on the screen. These are comparable measurements, but the automated statistical approach to measuring grain size inevitably introduces a fine bias in estimates of the population particle size statistics, by incorporating information from all of the objects in the image, therefore all the partial axes of some particles. The same test was adopted here, but simulated granular materials can provide a means by which to quantify the discrepancy between bulk particle size statistics and those derived from just the visible particles at the surface, which is the subject of section 3.2.

[10] Section 4 presents comparisons of observed versus estimated mean particle size and sorting, for material with (simulations only) and without (simulations, plus images of real sediment) an apparent void fraction. In section 5 we present an analytical solution for void to particle transition probabilities in a 2D section through a two-phase granular material which can be used to estimate the apparent void fraction using only the image's autocorrelation function, the plot of autocorrelation coefficient as a function of lag displacement, hereafter 'autocorrelogram'. The method is tested using 2D sections (slices) through some simulated granular materials, again generated using the methods in part 1. Finally, the uncertainties and implications of this work are discussed (section 6) before conclusions are drawn (section 7) .

## 2. Methods for Estimating Arithmetic Mean and Sorting

### 2.1. Stochastic Properties of an Image of Sediment

[11] *Buscombe et al.* [2010] exploited the stochastic properties of the image of sediment, as well as of objects created using a simple stochastic model of fragmentation, as revealed by the Fourier transform, to recover a geometric property —the mean of all intermediate ('b-axis') particle sizes —without having to measure the particles directly. Here, the same is achieved for the standard deviation of all intermediate particle sizes (sample arithmetic sorting).

[12] A demeaned image of granular material (where the mean intensity value is subtracted from each intensity value) composed entirely of touching particles at rest may be described as an homogeneous Gaussian random field, $I$, over two-dimensional spatial position $x$, $\{I(x = 1), I(x = 2),.., I(X)\}$. Such a field is completely described, in a statistical sense, by the power spectrum $\Psi(k)$, or equivalently its Fourier transform, the autocorrelation function $R(l)$ (over $l$ lags):

$$R(l) = \int_{-\infty}^{+\infty} \Psi(k) e^{-ikl} dl \tag{1}$$

and

$$\Psi(k) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} R(x) e^{-ikl} dl \tag{2}$$

where $k$ and $i$ are the wave number vector and imaginary unit, respectively. The centered, symmetrical 2D Fourier transform is the analytical equivalent to a diffraction pattern of a granular material body [*Preston and Davis*, 1976] without the phase problem of capturing a 2D representation of a 3D volume. This transform has enjoyed widespread use in the description [*Prince et al.*, 1995] and

reconstruction [*Liang et al.*, 1998] of granular material. Optical diffraction and Fourier methods have also been used to provide a metric of the size of non-touching, relatively low concentration samples or images of particles [*Gorecki*, 1989; *Momota et al.*, 1994]. For sediment grains without a preferred orientation (if the long axes of the visible parts of most grains are not pointing in the same direction), we can consider only the radial average of the (2D) power spectral density, where $\alpha$ are the angles over which the average occurs:

$$\Psi_r(k) = \int_0^{2\pi} \Psi(k)(k\cos\alpha, k\sin\alpha)\, kd\alpha. \tag{3}$$

[13] For sediment with a preferred orientation (e.g., imbricated gravels) the 2D autocorrelogram (equation (1)) must be used to find wave number k associated with that dominant angle, typified by the longest elongation of contours of autocorrelation [*Buscombe*, 2008]. The central limit theorem implies that a random process will be Gaussian, a distribution which arises whenever a large number of independent vectors are drawn from the same distribution [*Oppenheim et al.*, 1999]. An image of granular material satisfies this criterion because the image as a spatial Fourier decomposition contains Fourier coefficients derived from independent vectors which are themselves statistically independent. Such a condition arises because the image is an intensity map of numerous complicated objects with a full dynamic range of digitized pixel values, with no preferred orientation, no correlation between particle color and other physical characteristics such as shape and size (except from the case of heavy minerals which are commonly darker in color and smaller in size), and with no background intensity (even that arising from inter-particle shadows which tend to have a distinctive and relatively uniform intensity).

[14] Although an autocorrelogram contains no phase information, unlike the full 2D Fourier transform, it nevertheless provides statistical information about particle geometries. In principle, a complete statistical description of local maxima (individual particle centers within the image) can be extracted from the autocorrelation function. If the phase-angle distribution is random in the range [0, $2\pi$] it does not affect the autocorrelations and it becomes possible to generate Gaussian fields from the autocorrelogram using Fourier transforms [*Liang et al.*, 1998; *Koutsourelakis and Deodatis*, 2005]. Population statistics from the two-phase random field may therefore be obtained from its autocorrelogram despite its non-uniqueness and the loss of phase information, and without the requirement of first measuring the outlines of individual particles. This is possible because the Fourier transform of a Gaussian function of variance $\sigma_I^2$ is also a Gaussian shape but with variance $1/\sigma_I^2$:

$$I(x) = \frac{1}{\sigma_I\sqrt{2\pi}} e^{-x^2/2\sigma_I^2} \overset{\mathrm{F}}{\leftrightarrow} H(k) = e^{-\sigma_I^2/2} = e^{-k^2/2(1/\sigma_I^2)} \tag{4}$$

where $F$ denotes Fourier transform, or its inverse. Importantly, such a Gaussian random field is completely described, in a statistical sense, by its mean and autocorrelation function [*Ripley*, 1981].

### 2.2. Arithmetic Mean

[15] *Buscombe et al.* [2010] exploited these properties to show that an image of sediment can be modeled as a Gaussian random field and further that the autocorrelogram is related to the probability of finding two points separated by $l$ pixels and belonging to the same (void-particle) phase, and further to recover the mean of the size-density of particles, which is given as:

$$\mu = \int_S sP(s) = 2\pi/k_R \tag{5}$$

where $s$ are the $S$ particle sizes present with probability density $P(s)$, and $k_R$ (units of 1/pixels) is the wave number associated with the lag at which $R(l) = 0.5$. The main advantages of the method are: 1) the autocorrelogram, since it is drawn from the power spectrum, captures all scales of variability and therefore contains information on the entire range of wavelengths (particle sizes) present in the image; and 2) the mean is estimated directly without first determining the entire particle size distribution $P(s)$. Similarly, it should be possible to estimate $\sigma$ (the standard deviation of particle sizes) directly, again using information from $R(l)$ on the range of scales of objects (particles) within the image it represents.

### 2.3. Arithmetic Sorting

[16] The approach taken here to estimate heterogeneity in particle sizes is to quantify the divergence in form between the sample's autocorrelation $R(l)$ and that of an idealized material, hereafter termed $R_u(l)$, which has the same mean particle size but zero variance. This model for zero particle-size variance is nonlinear because the particles possess a distribution of shapes. For example, identical-sized tessellated squares would give a linear autocorrelogram of $-0.1l + 1$ [*Buscombe et al.*, 2010], but the theoretical form of cell distributions formed from other forms of tessellation are always nonlinear [e.g., *Muche and Stoyan*, 1992].

[17] Experimental evidence [*Rubin*, 2004; *Buscombe and Masselink*, 2009] and simulations [*Buscombe et al.*, 2010] have shown that the autocorrelogram for well-sorted granular material consist of a steep quasi-linear part between $0.5 \leq R \leq 1$, and a concave part between $0 \leq R \leq 0.5$. The shape becomes more exponential with smaller standard deviation of particle sizes. *Warrick et al.* [2009] noted that autocorrelograms associated with images of more poorly sorted sediment (a larger standard deviation of particle sizes) diverted more significantly from an exponential shape (usually either the linear part is shorter or the concave part flatter, or both).

[18] Based on the above observations it would therefore seem that a suitable model for $R_u(l)$ is a single exponential decay in autocorrelation (a principle also used in the field of photon correlation spectroscopy to estimate the particle size of a monodisperse particulate suspensions. See review by *Berne and Pecora* [2000]). The image's autocorrelogram is the sum of the exponential decays corresponding to each of the grain size fractions present. The form of such an autocorrelogram would become more

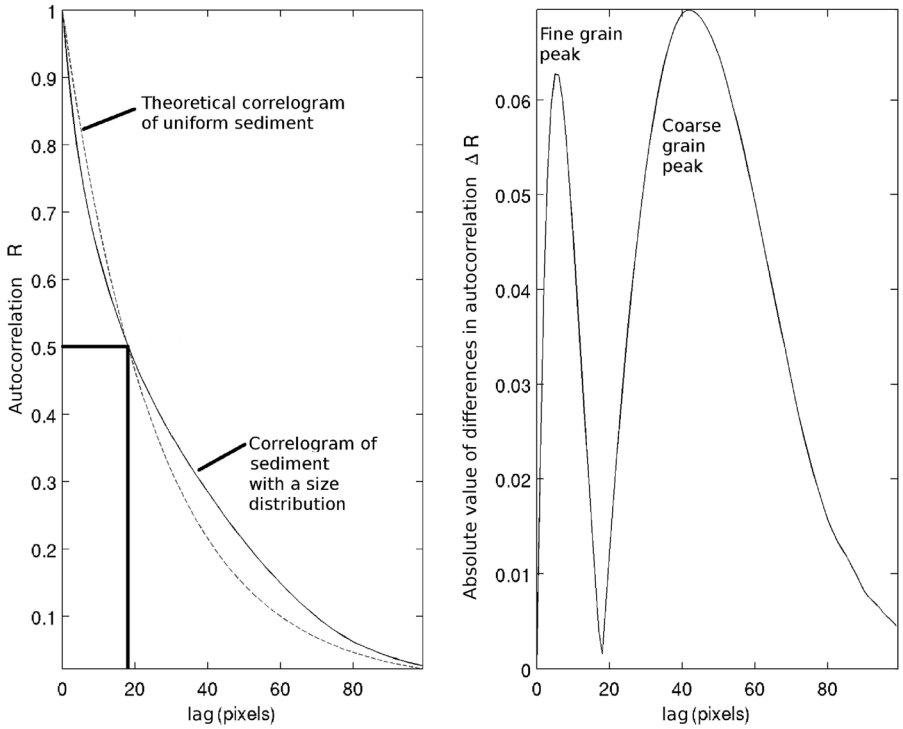


**Figure 1.** Schematic of the approach used to derive arithmetic sorting from an autocorrelogram. (left) The term $R_u(l)$ is used for a sediment with zero particle-size variance (solid) and empirical $R(l)$ is used for a sample with a distribution of particle sizes (dashed) but the same mean. (right) The differential typically shows two peaks associated with the fine and coarse tails, respectively, and a unique minimum at the lag associated with the mean.

dissimilar to an exponential shape with more grain size fractions present, consistent with the observations of *Warrick et al.* [2009]. It is therefore hypothesized that an exponential autocorrelogram shape represents a granular surface with given mean particle size and zero variance, and that exponential functions evaluated over $l$, scaled by $k_R$, tend to be suitable models for $R_u$.

[19] In order to remove the influence of any preferred orientation in the power spectrum, the 1D $R(l)$ is calculated from the 2D $R(l)$ (1) by radial-averaging [*Prince et al.*, 1995]. Hereafter, $R(l)$ refers only to the 1D, radially averaged autocorrelogram and $R_u(l)$ is therefore also a 1D function. Information at small lags is primarily on small wavelengths, but still may be affected to some degree by large wavelengths if the distribution is not perfectly symmetrical. Likewise, information at large lags depends primarily on large wavelengths. The central point (the lag at which $R = 0.5$) corresponds to the mean [*Buscombe et al.*, 2010]. The difference between empirical and idealized autocorrelograms is related to $\sigma$ because both have a unique minimum (i.e., zero) at the mean, so that the difference between the two curves is zero at that lag. This minimum of zero dictates that $\sigma$ is proportional to the integral of the entire difference in autocorrelograms, rather than two separate integrals for the portion of the autocorrelograms above and below the mean. Therefore, the divergence in form is quantified in terms of the absolute value of the difference between the two functions. Formally we may write the above as follows. The arithmetic standard deviation $\sigma$ is given by the definite integral of the difference between the sample's (empirically derived) autocorrelogram and that of an idealized material with same mean, $R_u$ but zero variance. Symbolically, treating $\sigma$ as a continuous random variable and the second central moment of $P(s)$:

$$\sigma = \sqrt{\int_S (s-\mu)^2 P(s)} = c \int_{L_0} [|R(l) - R_u| dl] \qquad (6)$$

where constant $c$ is chosen to scale the integral to be invariant of $\mu$. This constant is revisited later after suitable models for $R_u(l)$ are proposed. This new method is summarized graphically in Figure 1. The left panel depicts examples $R(l)$ (solid line) for an image of a granular surface (with a size distribution) and its corresponding $R_u(l)$ (dashed line). The standard deviation $\sigma$ is the product of $c$ and the definite integral evaluated over $l = [0, L_0]$ of the absolute value of the difference between $R(l)$ and $R_u(l)$ (right panel). $L_0$ is necessary because the autocorrelograms of random fields such as images of sediment tend to fluctuate around zero at long lag due to edge effects, nonuniform lighting and the presence of noise.

[20] Estimated autocorrelation coefficients become less reliable with increasing lags [*Bartlett*, 1946]. The suitable value for $L_0$ is suggested as the first lag where $R$ equals 0.

The explanation has a parallel in fluid mechanics, where the integral length scale of turbulence *Taylor* [1938] is calculated as the definite integral of the autocorrelogram, evaluated to the first lag where $R$ equals 0. This integral quantifies the spatial extent of correlations, and equals $2\pi$ multiplied by the spectral amplitude at wave number 0 [*Priestley*, 1981, p. 320].

[21] The Gaussian $e^{-k_R^2 l^2}$ [*Koutsourelakis and Deodatis*, 2005] is considered suitable as a simple model for $R_u$ because sediment images can be modeled as Gaussian random fields (section 2.1), and also because of this model's widespread use in modeling idealized random fields with zero variance [*Goff and Jordan*, 1988; *Holliger et al.*, 1993; *Fenton*, 1999; *Buscombe et al.*, 2010]; random point processes [*Davis*, 1986]; the statistics of fragmentation [*Grady*, 1990]; and because of its mathematical simplicity. Another important practical implication is that the constant $c$ in equation (6) can be determined using a quantity known as the scale of fluctuation, $\delta$, which quantifies the spatial extent of strong autocorrelations [*DeGroot and Baecher*, 1993; *Fenton*, 1999]. The Gaussian-exponential function has an analytical derivation for $\delta$, given by $1/k_R$ [*Uzielli et al.*, 2005]. Since $k_R = 2\pi/\mu$ (5), $c = \pi$ in order to scale the integral to be invariant of $\mu$.

[22] The use of the Gaussian for $R_u$ gives the following expression for $\sigma$:

$$\sigma = \pi \int_{L_0} \left[ |R(l) - e^{-k_R^2 l^2}| dl \right]. \tag{7}$$

Radial (1D) averages of the power spectral density of the sediment image reduce to, for a Gaussian autocorrelogram:

$$\Psi_r(k) = \frac{\sigma_I^2 L}{2} k l e^{-k^2 l^2/4}. \tag{8}$$

Other simple models such as exponential $e^{-k_R l}$ and cosine-exponential $e^{-k_R l \cos(k_R l)}$ [*Uzielli et al.*, 2005] also have analytical derivation for $\delta$ and could be used instead of the Gaussian. Competing simple models for $R_u$ related to the Gaussian-exponential, such as the Voronoi $(1 + 2k_R l)e^{-2k_R l}$ [*Muche and Stoyan*, 1992; *Buscombe et al.*, 2010] and von Karman [*Goff and Jordan*, 1988], are not considered further in this contribution because we do not know of an analytical derivation for $\delta$ [*Uzielli et al.*, 2005]. In other situations, models with no analytical solution to $\delta$ (hence $c$) may have applicability, in which case those parameters have to be determined empirically following one of the methods reviewed by *Uzielli et al.* [2005].

[23] The form of (6) applies when the image is composed wholly of particles, either fully exposed or partially hidden (in other words, where there is no apparent void fraction or 'porosity'). Such images include surfaces of unconsolidated granular materials imaged with a camera (see part 1 for a review). In instances where there is a void fraction, such as sections through granular material, (6) must be modified to account for the presence of said void fraction. The only modification required to (6) is that the integral is scaled to be invariant of $\mu$ using $(1 - \phi)c$, which has the limits $[0, c]$. Therefore, the relation for sorting accounting for known void fraction takes the form:

$$\sigma = c\phi \int_{L_0} [|R(l) - R_u| dl]. \tag{9}$$

[24] Similarly, the method of *Buscombe et al.* [2010] for calculating mean particle size is modified to account for the void fraction:

$$\mu = \phi 3\pi k_R. \tag{10}$$

Equations (5) through (10) give their results in dimensions of length in pixels. To convert to real units of length, outputs are multiplied by the spatial resolution of the image (in units of length/pixel). In the following, the Gaussian model (7) for $R_u$ has been adopted, both for images with and without an apparent void fraction.

## 3. Surfaces of Simulated Granular Material

### 3.1. Description

[25] The immediate application of the (3D) modeling approaches detailed in the companion paper (part I) has been to replicate a planform perspective of a sediment deposit, as would be captured by a photograph. The techniques have enabled a diverse set of simulations of granular materials with known properties and with which to evaluate the methods outlined in section 2. A few example images from various model outputs are shown in Figure 2. These are models of unconsolidated sediment surfaces with no apparent void fraction, with known population and apparent particle-size distributions. The images are composed of particles random in location, shaded, and 'packed' to a depth of several particles, so void space effectively disappears and the image looks like whole particles with parts of particles in the interstices. The reader is referred to part 1, a large proportion of which is devoted to the explanation of various spatial point-generating models which give rise to different particle packings. Here we have used model outputs from the following models, in order of decreasing regularity (determinism) in the packing of particles: 1) CVT-Halton (a centroidal Voronoi tessellation [*Du et al.*, 1999] based on particle centers generated by a Halton process [*Halton*, 1960]); 2) CVT-Uniform (a centroidal Voronoi tessellation based on a more uniform spatial distribution of particle centers); 3) NVT-Poisson (a normal Voronoi tessellation based on a Poisson (random) distribution of particle centers); 4) Strauss (one where a specified fraction of particle centers is allowed within a distance of any given particle center [*Ripley*, 1981]); 5) CP (Cluster Poisson: essentially a Possion distribution for particle centers with a built-in particle clustering mechanism) with small variance ($\sigma^2 = 0.01$; $\sigma$ is a parameter which controls the spread of particles in a cluster); and 6) CP with large variance ($\sigma^2 = 0.1$).

### 3.2. Surface-Subsurface Grain Size

[26] Granular simulations based on the NVT-Poisson model (defined above and described in part 1) have been used to quantify the effect of particle overlap on the mean size of particles apparent on the sediment surface. Models

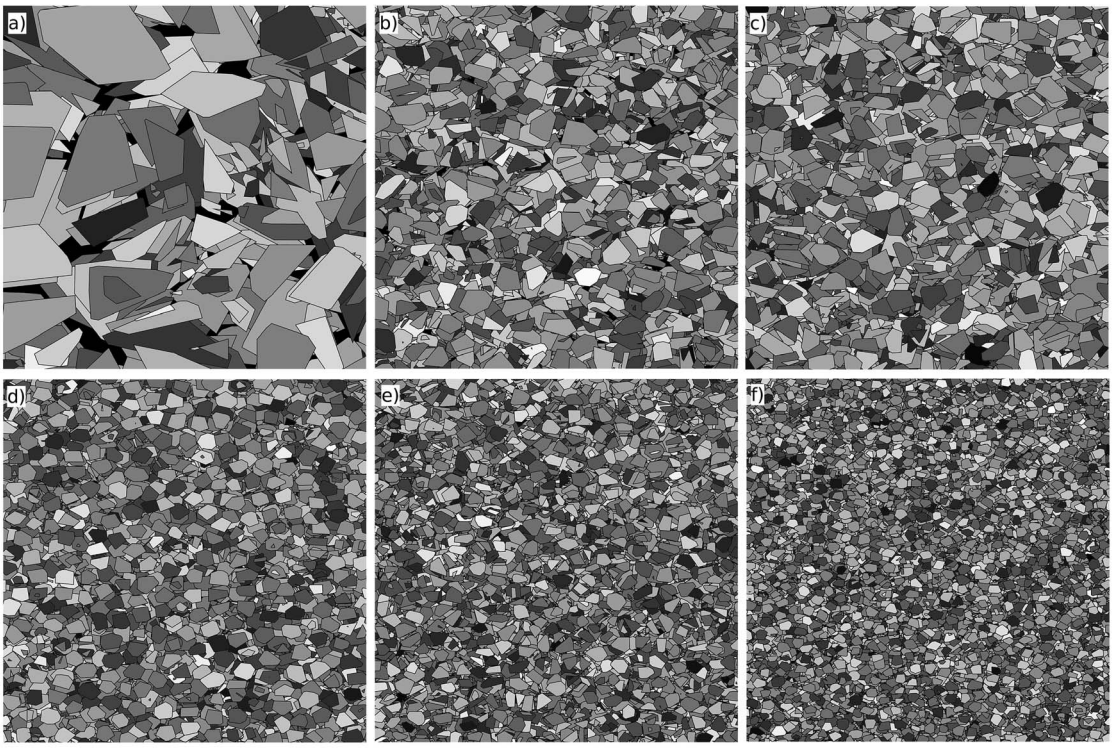


**Figure 2.** Projected images of the surfaces of a variety of synthetically generated 3D granular material. From left to right, and top to bottom, decreasing in mean particle size and increasing in the degree of deterministic structure to the particle packing. (top left) An inhomogeneous CP-based sediment with intensity $\lambda = 10$ and variance $\sigma^2 = 0.01$; (top middle) a CP-sediment with $\lambda = 10$ and $\sigma^2 = 0.1$; (top right) has been generated using a Strauss inhibition process to generate particle centers. The granular materials have been generated using (bottom left) NVT, (bottom middle) CVT-Uniform, and (bottom right) CVT-Halton spatial point processes.

are used instead of real samples because highly accurate measurements of particle size are required in very thin surface layers to a depth beyond which no particle is visible from the surface from an overhead perspective. Such measurements would be very difficult with real samples with the same accuracy as achieved using simulations (this is one of the uses for which the simulated-sediment models of part 1 were developed). The geometric projection of the surface (or any intersection) of a simulated 3D sediment volume onto the 2D plane which we call the 'image' of that surface has similar properties as images of real sediments, where parts of particles/cells are sitting on top of others, wholly or partially obscuring those underneath.

[27] A granular material composed of unconsolidated packed grains, at less than its angle of repose, has a bulk concentration (usually between 0.6 and 0.7). The top surface layer of grains, for example the grains you will pick up using a greased card, will be of a slightly lower concentration. This is because of relative grain-scale topography, caused by packing within the interstices of the grains below. The grains at the very surface have no grains above them to fill in their interstices, so the concentration is lower than that of the bulk. This only applies to surface grains. One can see the grains below this surface layer through their interstices.

[28] Particles on the surface of natural beds have been shown to have a statistically random spatial distribution of particle heights and locations [*Nikora et al.*, 1998] and the same is true of the elevations of particles in the surface layer of simulated granular materials. Figure 3 shows a surface of a 3D simulated sediment bed in which the particles are shaded according to their depth. Similar to surfaces of natural sediment beds, the 3D particles of the simulation touch and are packed randomly with an apparent pore space. In this regard model outputs are realistic, being how we perceive a plan view of a sediment surface.

[29] Volumes of granular materials were simulated with differing numbers of particles, with surfaces like those in Figure 2. The beds were between 15 and 40 particle diameters thick, depending on particle size, but were of identical volume. Sections were taken through each bed at a large number of discrete vertical coordinates through the volume. Figure 4 shows four such slices through a 3D bed. For each slice, the mean particle size was calculated. As one would expect, that size is always small where the tips of the particles are being intersected (Figures 4a and 4b). The depth to which particles may be viewed from directly overhead was found to be a consistent 4–5 mean particle diameters, and this trend was consistent across particle size and type of model used (note that for the construction of Figure 4, the CVT-Uniform model was used).

[30] The ratio of mean particle size per slice to mean size per volume (i.e., the 'intersected' mean, which is the mean of all sliced grain sizes), is shown in Figure 5 to be an increasing function of depth from the surface, expressed in

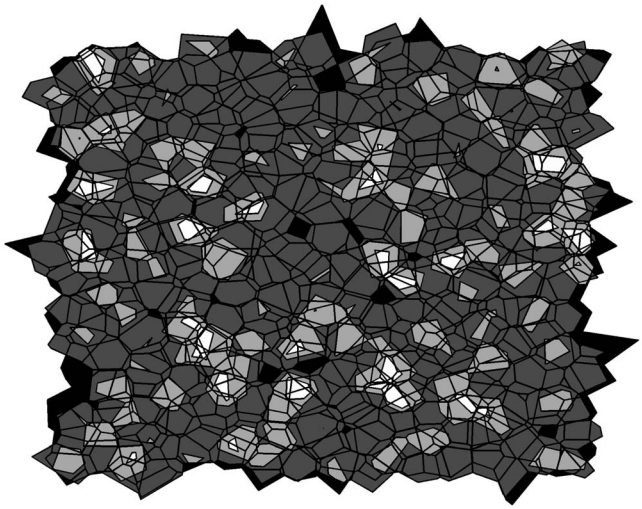

**Figure 3.** Four superimposed slices (sections) through the surface of a NVT-Poisson based 3D sediment bed, where the particles have been colored according to the depth of their center (white is 0 mean particle diameters, i.e., the very tip of the surface; light grey is 1; grey is 2; and black is 3). By 3 particle diameters depth, there tend to be no apparent void spaces (in other words, no gaps seen through the bed from overhead, the particles taking up all apparent space from this perspective), and beyond this particle size statistics of individual slices stabilize. Typically surface views of sediment contain particles whose centers are within 3 mean particle diameters of the tip of the sediment.

terms of the number of mean particle diameters depth. This can be explained by two phenomena which act in concert. The first is that the likelihood of slicing through the largest part of the grain will increase roughly toward the grain's center (of mass or otherwise defined). This not only affects the surface layer but also subsequent layers due to the influence of packing causing adjacent particles to have their centers at different relative elevations. The second process is that as the concentration increases with slice depth the likelihood of encountering a relatively large grain in the population increases. The increase in mean particle size as a function of depth is roughly exponential.

[31] Figure 5 shows that the point at which the rate of change of grain size with depth starts to plateau is also the approximate depth to which parts of subsurface grains are visible from the surface. This seems to be independent of average grain diameter. The implication for automated algorithms for sediment surface particle size is that such estimates (as well as the 'true' measurements against which we compare estimates, using point counting methods of apparent particles, be they whole or partially obscured) are likely to be smaller than the intersected mean as given by the mean of all intersected particle sizes in the volume, but that this effect is only relevant down to a depth of about 4 particle diameters.

## 4. Observations Versus Estimates

### 4.1. Simulated Surfaces

[32] The relationship for sorting (7) was tested using images of simulated granular material surfaces from the full range of model types described in section 3.1, a subset of

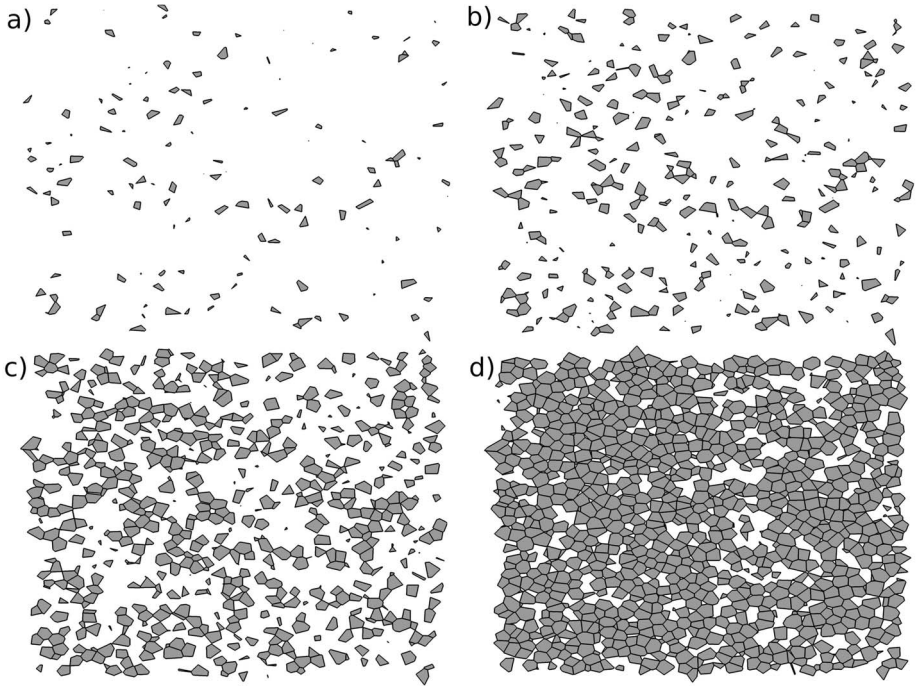


**Figure 4.** Slices through the surface layers of a NVT-Poisson based simulation of a sediment bed, similar to Figure 3 but in this case the slices (which are not identical) are separated: (a) top layer (1% of total depth of the bed, which is 40 mean particle diameters); (b) next layer down, about 1/2 particle diameter from top; (c) next layer down, about 1 particle diameter from top; and (d) about 2 diameters from top.

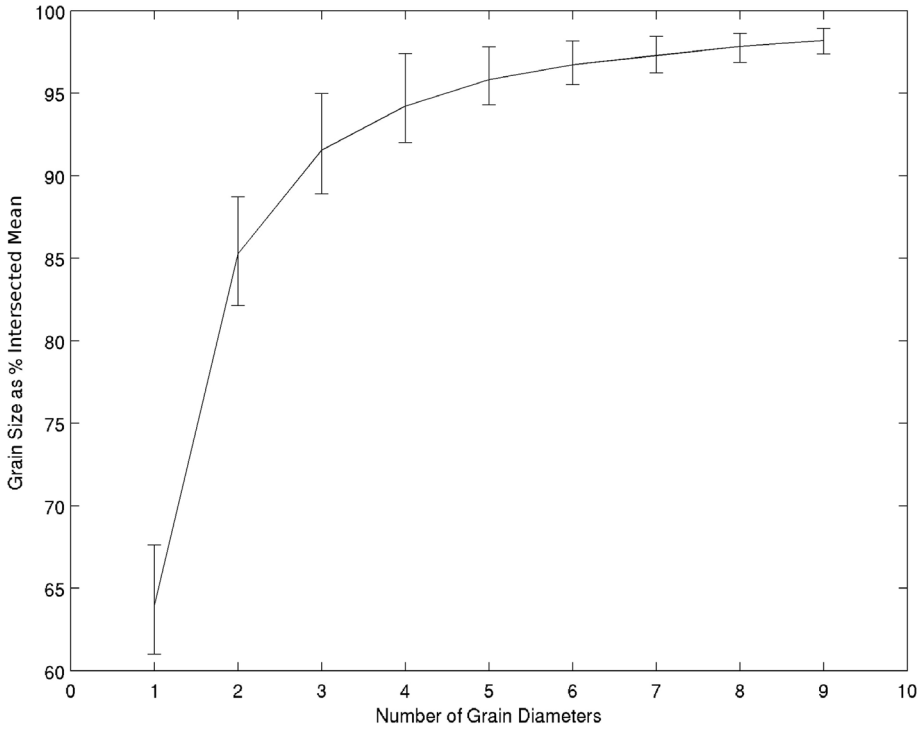


**Figure 5.** Mean particle size as a percentage of the mean of all intersected particles in the sediment (the same simulations as in Figure 4), calculated over a depth expressed in mean particle diameters. Error bars show the same statistic based on the maximum and minimum of intersected means. The mean of all particles from the surface to that depth is likely to be around 10% smaller than the intersected mean particle size due to the effects of partial particle concealment.

which are shown in Figure 2. The number of individual particles required for good estimates of the mean particle size using (5) was shown by *Buscombe et al.* [2010] to be more than approximately 1000. It was assumed that a similar number was required for estimates using (6), therefore images were generated containing between 1000 and 64,000 individual particles. For these simulated surfaces, the algorithm of *Buscombe et al.* [2010] predicts apparent mean particle size reasonably well: the errors are homoscedastic with an ensemble RMS error of 14.9% using (5) for mean size (Figure 6a). The ensemble RMS error for sorting is 13.9% using (7) (Figure 6b). Table 1 shows RMS errors for individual 'populations' of simulated material. Names for populations therein refer to the spatial model used in their construction (see section 3 for description and part 1 for a detailed explanation).

### 4.2. Images of Real Sediment Surfaces

[33] The sorting algorithm (7) was also tested using a set of 262 images from 8 different populations of natural unconsolidated granular surfaces (taken from beaches, rivers and continental shelves), both sand and gravel, taken with different camera/lighting systems. The images were mostly taken *in situ*, and often underwater, which required specialist hardware (the details of which are given by *Rubin et al.* [2007] and *Warrick et al.* [2009]). The set constitutes a subset of the populations used by *Buscombe et al.* [2010] for which a reliable estimate of sorting coefficient was available. This was achieved through manually measuring the intermediate axis of every whole particle in each image (see *Barnard et al.* [2007] for a more complete description). This is a time-consuming process, but the only way to independently measure the same particles as the estimation technique, thereby achieving a truly comparable metric of particle size [*Warrick et al.*, 2009; *Buscombe et al.*, 2010].

[34] The ensemble RMS error in mean size of 30.9 % (Figure 7, circles) using (7) (the Gaussian model), which reduces to 27.4 % when corrected for bias (Figure 7, diamonds). Bias-correction is a slope correction applied to estimates, found as the matrix division of estimates and observations (which forces the intercept through the origin). This calibrates for population-specific factors which may be present in, or apply to, images of natural particles. These factors are discussed in the next section. Table 2 lists uncorrected and corrected RMS errors, and the slope correction, for each of the 8 populations.

### 4.3. Simulated Granular Material With Void Fraction

[35] Simulations of sediment 'sections' (inherently 3D granular material simulated using the CVT-Halton methods outlined in part 1), with apparent void fraction ($\phi$) owing to the sectioning effect (examples shown in Figure 8) were used to evaluate (9) and (10) for, respectively, sorting and mean particle size when apparent void fraction is known. In these simulated sections through 3D granular material, particles have not been allowed to touch. Instead, particle locations have been held constant and the void fractions have been increased to varying degrees. Therefore, we are

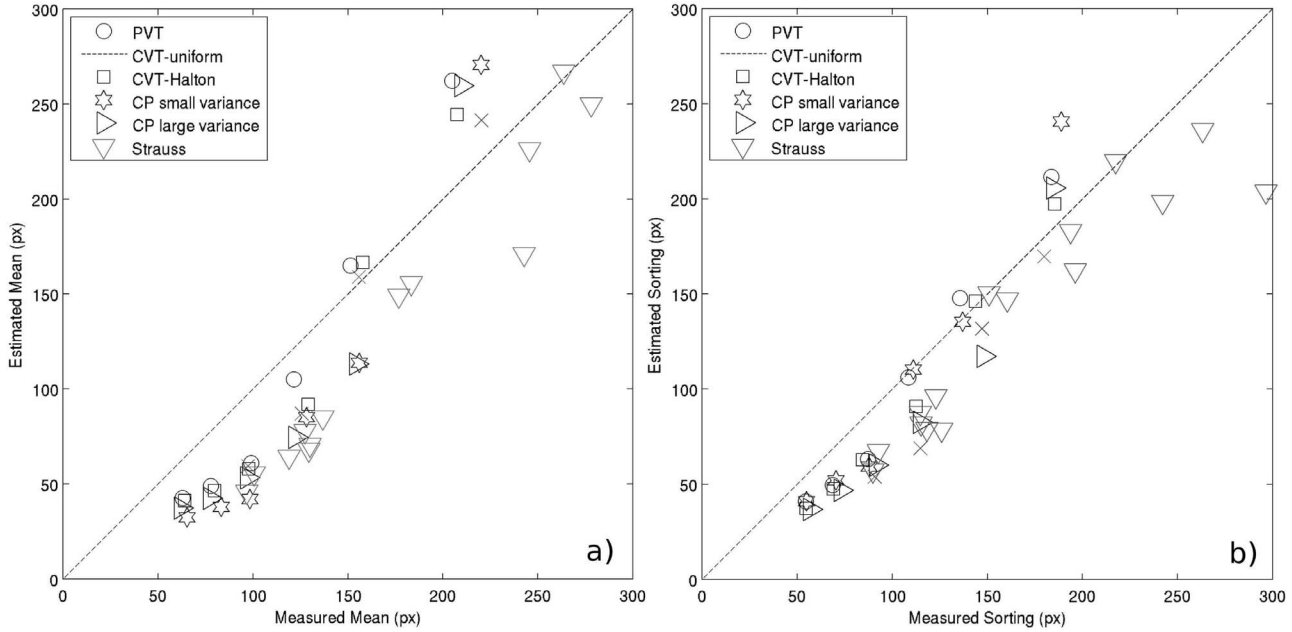


**Figure 6.** Estimated versus (a) true mean and (b) arithmetic sorting for 44 simulated sediment surfaces from a range of models (detailed in section 3.1) similar to those depicted in Figure 2. Ensemble RMS errors are 14.86% and 13.9%, respectively. The 1:1 relationship (dashed line) is also shown. All units are pixels.

testing only the influence of the basic structure of the bed and porosity, not factors such as grain coalescence. Using these beds, the ensemble RMS error between sorting observations and estimates is 10.4% using (9) using $R_u$ and $c$ appropriate for the Gaussian model (Figure 9a). Sorting estimated using (9) is very sensitive to correct measurement of void fraction. Here, the apparent (2D) void fraction is known precisely, but in images of real sectioned consolidated sediment it would be necessary to use careful thresholding techniques to measure void fraction exactly.

[36] The use of (10) to estimate mean particle size is not as good as estimates of sorting using (9), with an ensemble RMS error of 36% (Figure 9c, circles). It was found that error goes with $\phi$ such that, where $O$ and $E$ are observations and estimates:

$$\frac{O-E}{O} = 7.6\phi^3 - 16.6\phi^2 + 12.3\phi - \pi. \qquad (11)$$

(Figure 9b) using a cubic polynomial least squares fit with the constraint that the curve passes through $\phi = 0$ where $(O - E)/O = \pi$ (see Appendix A). However, the corrected estimate of the mean, which reduces the RMS error to 9.2%, requires $O$. Therefore in practice error must be evaluated as $O - E$ which, using the method outlined in Appendix A, requires estimating what value $(O - E)$ takes when $\phi = 0$. If this is estimated visually to be $-180$ $(-\pi)$, the correction of (10) takes the form:

$$\mu = \phi 3\pi k_R + (393\phi^3 - 880\phi^2 + 675\phi - 180) \qquad (12)$$

which reduces the ensemble RMS error to 8.9%. The corrected estimates of the mean particle size are shown as stars in Figure 9c. This method is only applicable if void fraction is known or well approximated.

### 4.4. Real Granular Material With Void Fraction

[37] The methods for sorting and size given by equations (9) and (10) were tested using an image of thin section through a sample of Folkestone sandstone [*Cresswell and Barton*, 2003]. The image is of poorly cemented quartz sands (Figure 10a reproduced from *Cresswell and Barton* [2003, Figure 2]) with an apparent void fraction of $\phi$ = 0.4. The standard deviation and mean of particle size calculated using (9) and (10), respectively, compared very well with the respective measures made by manually counting the on-screen intermediate diameters of particles (Figures 10b and 10c). Measured $\sigma$ and $\mu$ of 53.7 and 130.7 pixels, respectively, compared with estimates of 54.7 and 127.9 pixels. However, results are very sensitive to specification of $\phi$ (for example, using 0.35 increased percentage error in estimates of the mean from approximately 2% to 14%).

## 5. Void-Particle Transition Likelihoods and First-Order Approximation for Void Fraction

[38] In this section, we present a theoretical derivation of the first-order Markovian properties of a random binary

**Table 1.** Summary of RMS Errors (%) for Sorting and Mean Estimates Based on Tests With Simulated Granular Surfaces

| Population | Mean RMS Error (%) | Sorting RMS Error (%) |
|---|---|---|
| NVT-Poisson | 31.6 | 22.16 |
| CVT-Uniform | 31.5 | 17.2 |
| CVT-Halton | 9.42 | 8.84 |
| CP small variance | 20.8 | 25.7 |
| CP large variance | 31.9 | 15.27 |
| Strauss | 25.2 | 22.1 |
| Ensemble | 31.36 | 24.43 |

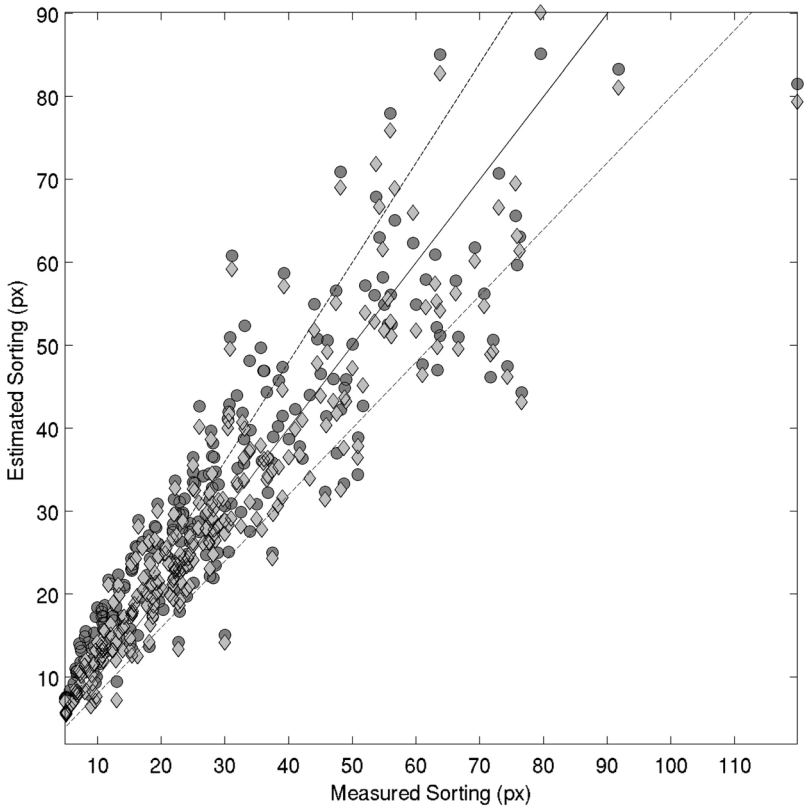


**Figure 7.** Measured arithmetic sorting versus that estimated versus using equation (7) (filled circles) and bias-corrected estimated sorting (diamonds). The slope adjustment is given in Table 2. The 1:1 relationship (solid line) and ±20% (dashed lines) are also shown. All units are pixels.

(void, particle) field constructed from the Voronoi tessellation of a simple Poisson (random) point process. Given the utility of such a tessellation for the modeling of granular material, as described at length in part 1, we argue that this can be used to estimate the one-step pore-particle transition probability matrix of a 2D section through a volume of granular material. One practical corollary is that the 1-step pore-particle transition probability matrix can be expressed in terms of the (easily computed) empirical autocorrelogram if the void fraction is known. This derivation is also shown.

[39] The granular section may be considered in continuous space and two intermittent phases ($q$); void (1; $v$) and particle (2; $p$). The space over which each phase exists follows an exponential distribution $P[q\ ;\ \Delta x] = e^{-(\Delta x/X_q)}$ where $\Delta x$ is a spatial step (in any dimension), and $\overline{X_q}$ is the mean distance over which state $q$ persists [*Cressie*, 1993].

[40] Assuming no sub-spatial step variation, the matrix of single-step transition probabilities of phases $i$ and $j$ can be expressed:

$$P_{ij} = \begin{pmatrix} P(v|v) & P(v|p) \\ P(p|v) & P(p|p) \end{pmatrix} = \begin{pmatrix} e^{-\Delta x/\overline{X_1}} & 1 - e^{-\Delta x/\overline{X_1}} \\ 1 - e^{-\Delta x/\overline{X_2}} & e^{-\Delta x/\overline{X_2}} \end{pmatrix} \quad (13)$$

**Table 2.** Summary of Tested Populations and Errors Associated With Application of Equation (6) for Arithmetic Sorting[a]

| Population | N | Sediment | RMS Error, Uncorrected (%) | RMS Error, Corrected (%) | Bias Slope |
|---|---|---|---|---|---|
| Santa Cruz 1 | 23 | Silt/Sand | 18.22 | 13.89 | 1.12 |
| Elwha | 55 | Sand/Gravel | 17.51 | 16.17 | 0.94 |
| Pescadero 1 | 55 | Sand | 40.08 | 16.75 | 1.11 |
| Colorado River | 13 | Sand | 17.42 | 16.95 | 1.04 |
| Santa Cruz 2 | 13 | Sand | 27.64 | 18.35 | 0.83 |
| Quarry | 15 | Gravel | 20.03 | 19.26 | 1.05 |
| Pescadero 2 | 35 | Sand | 43.49 | 25.78 | 0.77 |
| Slapton | 53 | Gravel | 35.25 | 35.12 | 0.97 |
| Total or Ensemble Mean | 262[b] | | 30.85[c] | 27.44[c] | |

[a]N refers to the sample size (number of images of sediment) in each population;bias slope is the correction required for the measured-estimated relationship to pass through the origin (referred to elsewhere in the table as 'corrected').
[b]Total.
[c]Ensemble mean.

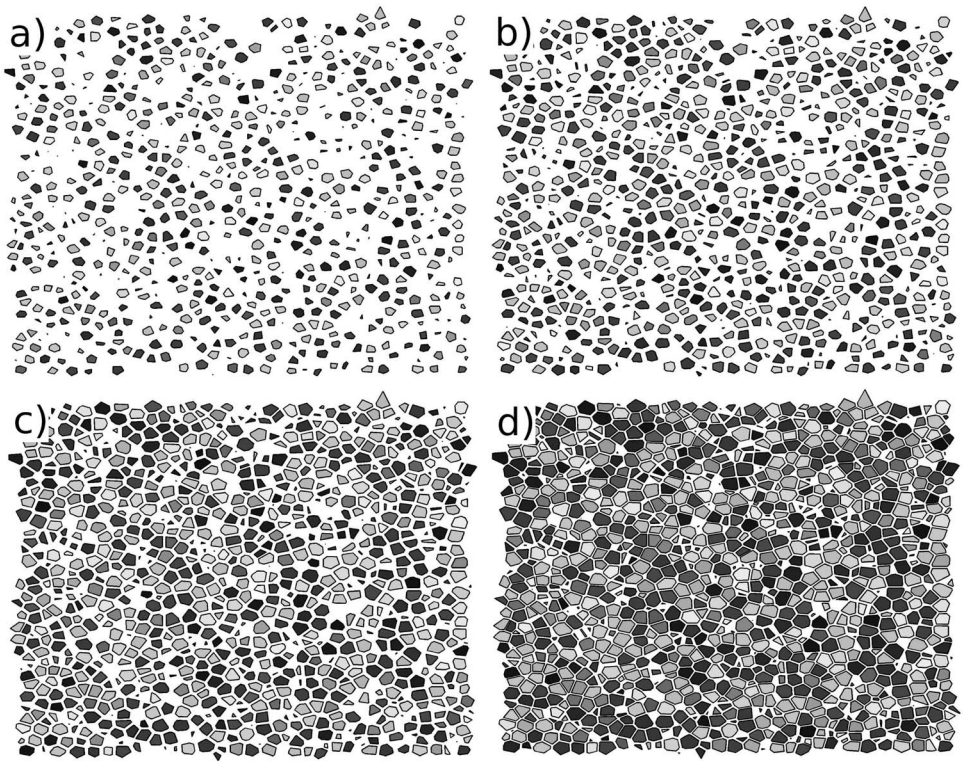


**Figure 8.** Example sections through CVT-Halton based simulated granular material with apparent void fraction: (a) 248 particles (mean size = 44 pixels), $\phi = 0.79$; (b) 314 particles (mean size = 49 pixels), $\phi = 0.67$; (c) 367 particles (mean size = 56 pixels), $\phi = 0.5$; (d) 431 particles (mean size = 61 pixels), $\phi = 0.29$.

where $\overline{X_1}$ is the mean distance as pore and $\overline{X_2}$ is the mean distance as particle. Noting the limits

$$P'_{ij} = \lim_{\Delta x \to 0} \frac{P_{ij}(x + \Delta x) - P_{ij}(x)}{\Delta x} \tag{14}$$

$$\lim_{x \to 0} P_{ij}(x) = P_j \tag{15}$$

$$\lim_{x \to 0} P'_{ij}(x) = P_j \tag{16}$$

and defining $P_1$ is the probability of a void and $P_2$ the probability of a particle, where $P(p)$ as the probability that a point lies on a particle, the matrix of single-step transition probabilities can be rewritten as:

$$\frac{P_1}{\overline{X_1}} \frac{P_2}{\overline{X_2}} = 0 (x \to \infty) \tag{17}$$

$$P_1 = \frac{\overline{X_1}}{\overline{X_1} + \overline{X_2}} \tag{18}$$

$$P_2 = \frac{\overline{X_2}}{\overline{X_1} + \overline{X_2}} \tag{19}$$

$$P_2 = P(p). \tag{20}$$

[41] $P_{ij}$ has a 'steady state' marginal probability distribution which will satisfy [*Kemeny and Snell*, 1960]:

$$\Pi = \lim_{x \to \infty} (P_{ij}) = P_j = P_{ij}\Pi, \quad \Pi \geq 0, \quad \sum \Pi = 1. \tag{21}$$

The ratio of the joint probabilities of $I_i$ and $I_{i-l}$ and marginal probabilities of $I_i$ gives their conditional probabilities $P(I_{i-l}|I_i)$ (which is termed the 'autorun' function by *Sen* [1978, 1984]) and can be expressed as:

$$P(I_{i-l}|I_i) = \frac{P(I_{i-l}, I_i)}{P(I_i)}. \tag{22}$$

[42] The information necessary to statistically construct $P(I_{i-l}|I_i)$ is given by the lag $l$ (spatial) autocorrelation estimate of $\{I(x = 1), I(x = 2),.., I(X)\}$ observations, $R(l)$. Inserting (22) into (13), $P_{ij}$ may be rewritten as:

$$P_{ij} = \begin{pmatrix} P(I_{i-1}|I_i) & \frac{\phi}{1-\phi}(1 - P(I_{i-1}|I_i)) \\ 1 - P(I_{i-1}|I_i) & 1 - \frac{\phi}{1-\phi}(1 - P(I_{i-1}|I_i)) \end{pmatrix}. \tag{23}$$

[43] The conditional probability $P(I_{i-1}|I_i)$ will hereafter simply referred to as $r$. In the absence of a known joint probability distribution, following *Sen* [1978] for the relationship between $R$ and $r$ (assuming $I$ is Gaussian distributed

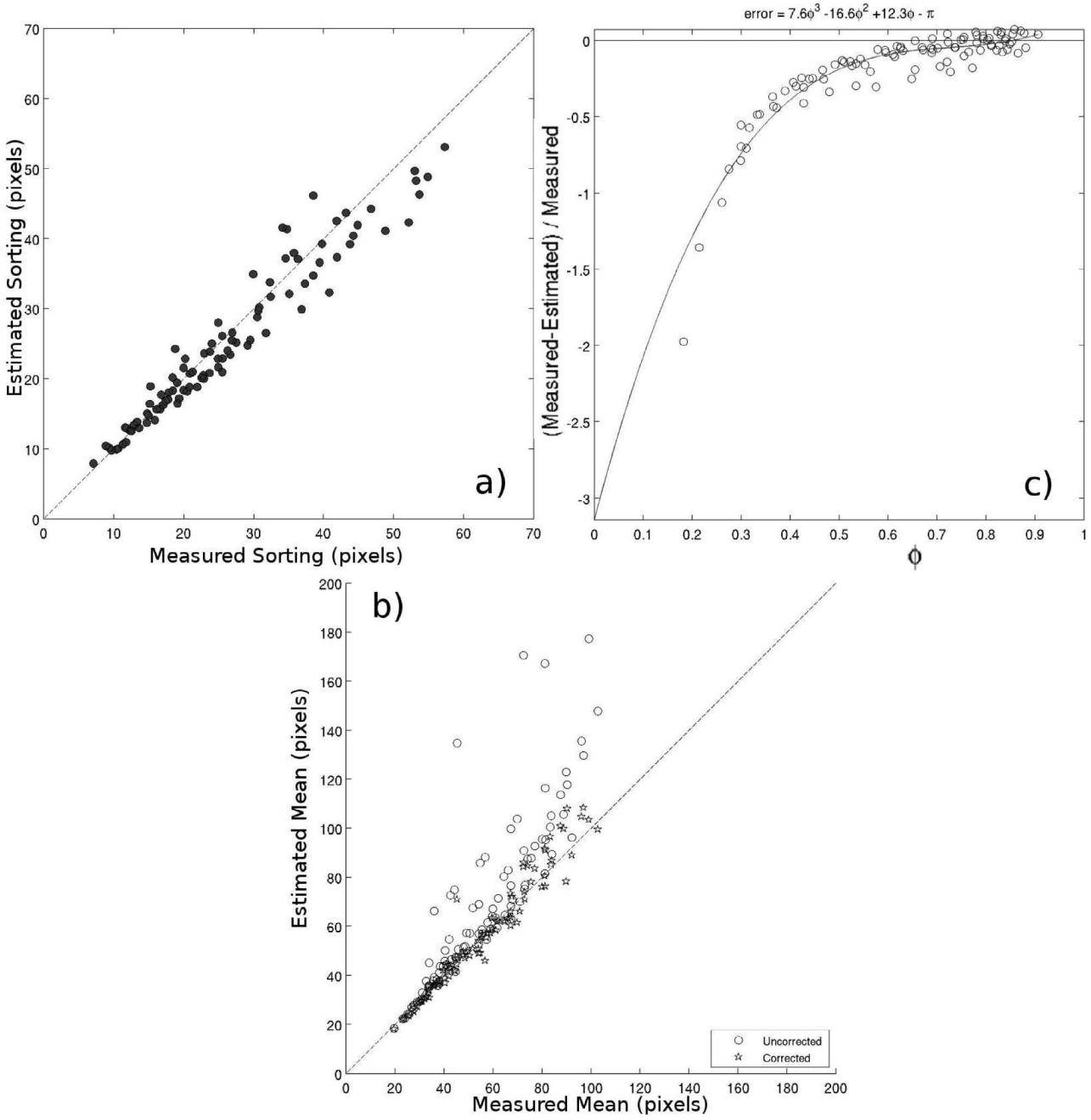


**Figure 9.** (a) Estimated versus true arithmetic sorting for 114 sections through simulated granular material, with ensemble RMS error of 10.43% (dashed line is the 1:1 relationship); (b) Estimated versus true arithmetic mean size, with ensemble RMS error of 36% for uncorrected estimates (equation (10), circles; dashed line is the 1:1 relationship) and 9% for estimates corrected using the function in (12); (c) normalized error as a function of void fraction $\phi$, with fitted function according to the method in Appendix A.

with respect to $x$), (23) can be rewritten, for known $\phi$ and $R_1$ ($l = 1$), as:

$$P_{ij} = \begin{pmatrix} \epsilon & \frac{\phi}{1-\phi}(1-\epsilon) \\ 1-\epsilon & 1-\frac{\phi}{1-\phi}(1-\epsilon) \end{pmatrix}. \tag{24}$$

where $\epsilon = [1/2 + \pi^{-1} arcsin(R_1)]$. For this binary case (21) is solved as the following matrix division:

$$\Pi = \frac{\begin{pmatrix} 0 & 1 \end{pmatrix}}{\begin{pmatrix} P(p|p) - 1 & 1 \\ P(g|p) & 1 \end{pmatrix}} \tag{25}$$

and the solution reduces to

$$\Pi = [\phi, 1-\phi]. \tag{26}$$

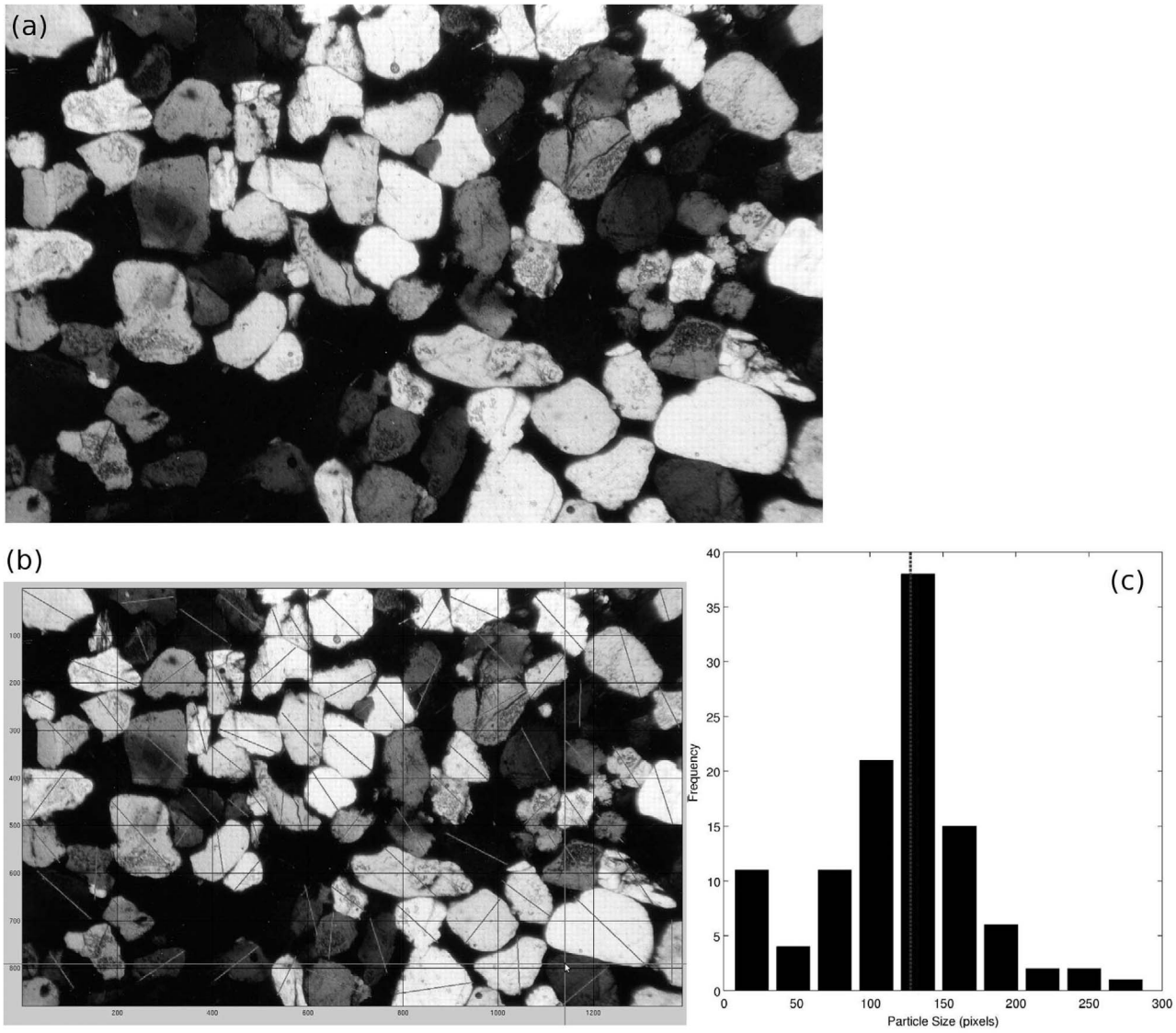


**Figure 10.** (a) Thin section of a poorly cemented, quartzose fine, medium and coarse sand (Folkestone formation, southeast England) imaged under crossed polar light. The field of view is 3mm across. Detrital quartz grains are white, grey and black, and the pore spaces black. This is sample A2 reproduced from *Cresswell and Barton* [2003, Figure 2] with an apparent porosity of 0.4. Image is 879 × 1386 pixels; (b) intermediate diameters counted manually on-screen; (c) histogram of manual point counts of particle size. Dashed vertical line indicate the mean grain size estimated by application of equation (10).

[44] Using this theory, the probabilities of pore-particle transition may be estimated from the (very easily computed) autocorrelogram of the granular material. Transition probabilities were measured for the cemented sand sample of Figure 10a, binarized so particles are 1 and voids are 0, as shown in Figure 11. Transition probabilities estimated using (24) from the image's autocorrelogram are not exact (Table 3) but are well approximated to first-order and good enough to predict $\phi$ calculated using (26). In this example a measured $\phi = 1 - (\sum I_x/N) = 0.4081$ compared to an estimated $\phi = 0.4029$, which further supports the basic theory presented above. Similar tests were conducted using sections of simulated granular material (using the CVT-Halton model; see section 3.1 for definition) with a range of porosities and number of particles. The transition probabilities based on autocorrelation (25) are shown as lines in Figure 12 and the actual values as symbols, showing that the estimates are only approximate.

## 6. Discussion

### 6.1. Sedimentological Basis for $R_u$

[45] The new techniques for estimating sorting given by (6) and (9) outlined here are consistent with the observations of *Warrick et al.* [2009] who noted that correlograms associated with images of more poorly sorted sediment (a larger standard deviation of particle sizes) diverged more significantly from an exponential shape. That shape we now hypothesize to represent a granular surface with given mean particle size and zero variance.

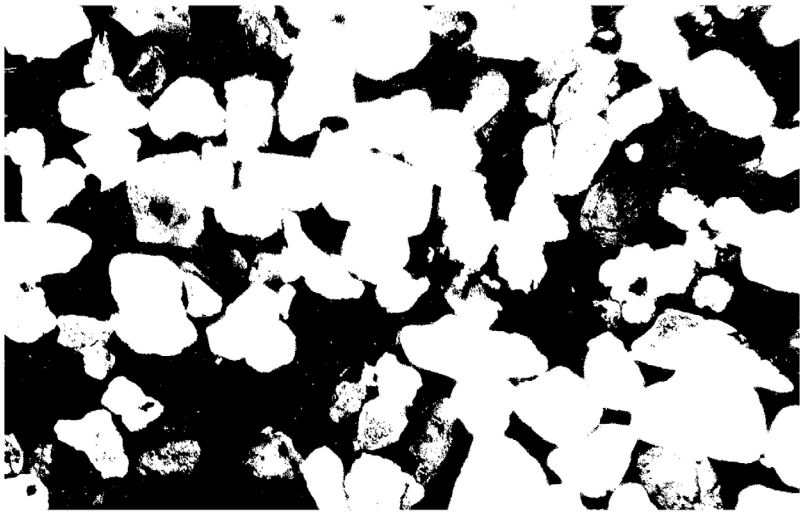

**Figure 11.** Binarized version of the cemented sand section shown in Figure 10a, with apparent porosity $\phi = 0.4$.

[46] Since the autocorrelogram can be derived directly from the image, and the mean can be calculated using the method of *Buscombe et al.* [2010], the method for calculating $\sigma$ has two points of adjustment, which are the equation for $R_u(l)$, and defining a suitable lag for $L_0$. This can be regarded as advantageous considering the multifarious structural nature of sediment, and the variety of devices through which granular material may be imaged with different spatial resolutions. However, it also means that estimation of arithmetic sorting is not as straight-forward in application as the method proposed by *Buscombe et al.* [2010] for mean particle size, because one must choose on a model of idealized material with zero variance, $R_u$. Here we considered a simple model —the Gaussian-exponential (7) —with which to evaluate against observations. It is possible, however, that the model-observation mismatch is minimized by a model for $R_u$ not considered here, and further that there is an optimal model for each population granular material under scrutiny. Further work is needed to ascertain whether there is a sedimentological rather than purely statistical basis for choice of model for $R_u$.

### 6.2. Evaluation

[47] It has been shown here that calibration is often necessary, with images of real sediment surfaces, to reduce the RMS error to an acceptable level by removing the population-specific bias due to a number of potential factors (Figures 7 and 9). These factors are inherently more difficult to model and could be potential sources of error in the application of the technique. The two main sources of error are: 1) the estimation of the mean particle size (algorithm error) and; 2) the sedimentological structure of the bed. Since the former and the latter are also related, it would be very difficult to separate their relative contributions to the total error.

[48] Examples of factors related to the sedimentological structure of the bed include the influence of particles so small they are sub-pixel in diameter, or so large they occupy a significant proportion of the image. In both instances, the estimated particle size statistics would be weighted with a coarse bias. Other potential problems could be distributions with multiple peaks, granular coatings (e.g., biogenic) and intragranular variability (pock-marks, scratches, and other shading-variations within individual particles). This latter factor was found by *Buscombe et al.* [2010] to cause very small errors ($\leq$10%) in estimated means, but strong reflections from crystal facets, where present, could constitute a significant source of error.

[49] Experience dictates that all of these are likely, and at least one of these factors is highly probable in most images of natural particles, especially those taken in the field. It may therefore be unreasonable to expect the methods outlined in this contribution to be applicable to all images of real sediment without significant error. In this instance, the best case scenario is that the source(s) of error(s) can be removed by filtering, or are site-specific and can be corrected by calibration. The fact that the errors are minimized to a satisfactory level using a calibration taking the simplest possible form (i.e., slope correction) means errors are specific to, and vary with, sediment populations rather than individual sample images ('correcting' each sample individually would of course render the 'automated' technique useless).

[50] Like the method of *Buscombe et al.* [2010] for estimation of the mean from the frequency transform of an image of sediment, the techniques described here for estimation of arithmetic standard deviation of particle sizes do not require resolution of the entire particle size distribution.

**Table 3.** Pore-Particle Transition Probability Matrix (13) Containing Probabilities Estimated Using Equation (24) Inside Parentheses and Those Measured From the Image are Outside Parentheses[a]

| Transition | $v$ | $p$ |
|---|---|---|
| $v$ | 0.9663 (0.9095) | 0.0337 (0.0905) |
| $p$ | 0.228 (0.0603) | 0.9772 (0.9397) |

[a]The term $v$ refers to voids and $p$ is particle such as in equation (13), $\rightarrow$ represents transition.

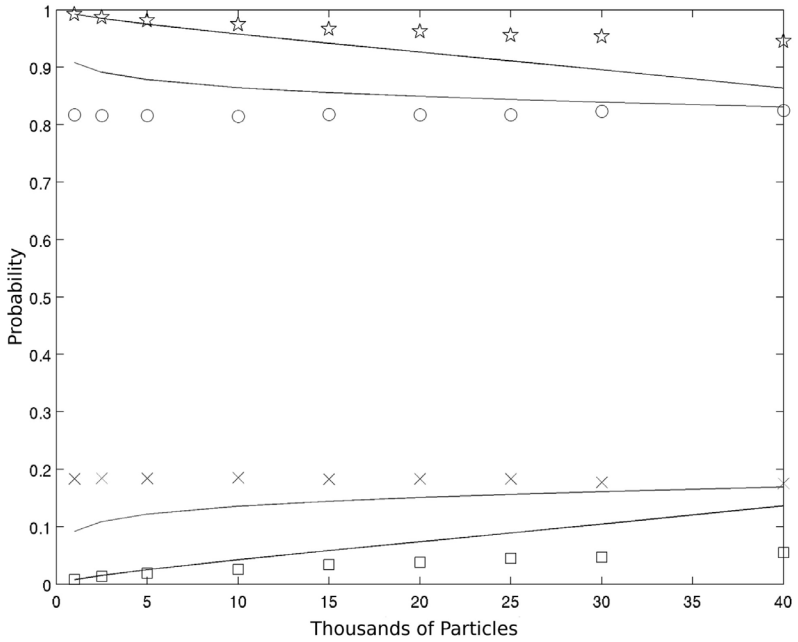


**Figure 12.** Stochastic statistics of a set of 2D sections through 3D homogeneous granular material, simulated using the CVT-Halton method described in part 1. Transition probabilities measured directly from the image (symbols) and estimated using (26) (lines). The symbols are as follows: stars = $P(p|p)$ (transition particle pixel to particle pixel), squares = $P(v|v)$ (transition void to void), circles = $P(p|v)$ (particle to void) and crosses = $P(v|p)$ (void to particle).

We therefore call these *direct* estimates. The disadvantage of such an approach is that it is not possible to estimate given percentiles of the particle-size distribution (widely used in geomorphology for example to quantify roughness) without first modeling the distribution, given the mean and standard deviation, using some assumed distribution shape (normal, lognormal, etc).

[51] A variety of quantitative measures are used to describe spread of a distribution of particle sizes, but none are universally accepted [*Fieller et al.*, 1992]. The arithmetic standard deviation, in linear units (rather than the sedimentologic phi transformation), seems an obvious first choice for a measure of particle-size distribution spread directly from the Fourier transform of the image of granular material, it being the 2nd central moment (or moment about the mean) of the distribution. It follows from intensities in images of sediment following the central limit theorem (4) that the most analytically tractable direct estimate of sorting is an arithmetic standard deviation. Other advantages of this metric include the use of information from the whole distribution (whereas graphical statistics, for example, use only certain percentiles of the cumulative size-distribution) and, using linear units, errors are relatively easy to compute, and the result is more readily understood and utilized. This latter point is perhaps especially important if this work finds potential use in fields outside of sedimentology/geology in the wider geophysics community as well as in engineering and industry. For example, industrial applications might require a non-intrusive, unsupervised measure of the spread of the distribution of granular sizes for real-time quality control.

[52] Other metrics of the spread of particle size distributions such as the geometric standard deviation, given by (using the same notation as (6), where log denotes natural logarithm):

$$\sigma_g = e^{\sqrt{\int_S (\log s - \log \mu)^2 P(s)}} \tag{27}$$

are more common in hydrology and geomorphology because bulk particle-size distributions of natural sediments are typically observed and assumed to be lognormal in distribution. In lieu of a direct measure of $\sigma_g$ like (6) for arithmetic $\sigma$, direct mappings between the measures are presently only achieved through empirical means for a given sedimentological population. For example, for a population of real images of sediment (population 5 in Table 2 was chosen for illustrative purposes because it has an RMS error between observed and estimated arithmetic sorting coefficients which corresponds approximately to the average of all samples) the observed $\sigma_g$ versus estimated $\sigma$ (Figure 13b) can be empirically corrected with a linear least squares fit for direct comparison of the two metrics (Figure 13c).

[53] The analysis of section 3.2 demonstrated that particles, either whole or in part, are visible from an overhead perspective to a depth of approximately 4 mean particle diameters. This is only the first step in the quantification of apparent versus true particle size distribution (in other words, the effect of particle overlap in the planform perspective of a sediment creating a visible particle size distribution which is different from the true distribution of the particle population). To do so would require an algorithm to 1) identify all of partial grains as viewed from the surface; 2) measure both their apparent and true axes in the orientation in which they are viewed; 3) use this information to quantify the effects on the algorithms for particle size which use only

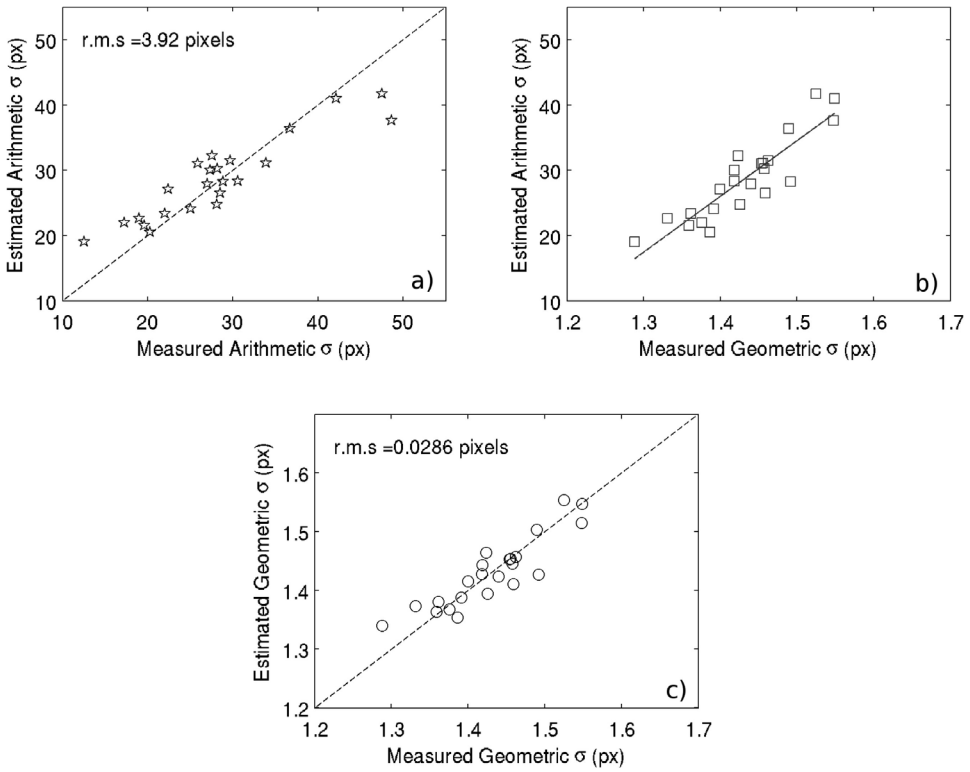


**Figure 13.** Comparison between arithmetic and geometric sorting coefficients for one set of images of real sediment (population 5 in Table 2): (a) measured versus estimated arithmetic sorting using (7); (b) measured geometric (27) versus estimated arithmetic sorting using (7); and (c) measured geometric sorting (27) versus estimated geometric sorting using (7) adjusted using a linear least squares fit (slope = 0.0094, and intercept = 1.1612).

the statistical information as contained in the autocorrelogram. For the latter part especially, there remains much to do.

### 6.3. Application

[54] The new expression for mean particle size (10) and sorting (9) in material with an apparent void fraction has performed reasonably well with sections through simulated granular material. These are images of irregularly shaped and dilutely distributed particles in a volume, achieved by using a random spatial point distribution such as the Poisson distribution. It thus appears to be a promising advance on similar techniques with a spherical-particle assumption [e.g., *Gorecki*, 1989; *Momota et al.*, 1994]. While in many practical applications of the technique it is disadvantageous that void fraction must be known in advance, we have shown that errors scale with void fraction, and that errors are small when void fraction is greater than 50% (Figure 9c). As such it may complement automated sizing of particles, as such the latest methods in thin section microscopy [e.g., *Torabi et al.*, 2008] for 2D images of particles, and in-line particle holography [e.g., *Graham and Nimmo-Smith*, 2010] or other optical techniques based on scattering theories [e.g., *Pereira and Gharib*, 2004] for 3D images of particles.

[55] The new method in section 5 for the estimation of $\phi$ from a two-phase (particle, void) image using only information contained in the autocorrelogram has limited practical utility in its current form because a) thresholding is required to binarize the image, and b) calculation of areal void fraction from a binary image is trivial. However, it is at least of statistical interest, plus we foresee no theoretical obstacle to it serving as a simple starting point for estimation of $\phi$ directly from non-thresholded images, which would have considerable utility.

## 7. Conclusion

[56] New formulas have been proposed for the arithmetic sorting coefficient of granular material with and without apparent void fraction. The measure is equivalent to the standard deviation of apparent intermediate particle diameters. The method uses the autocorrelogram of the sediment image, estimated from its Fourier transformation, and an autocorrelogram model for a material with the same mean particle size but with zero variance in particle size (i.e., a uniform particle size distribution). A number of available options for such a statistical model provide inherent flexibility in the approach. In addition, a modification to the algorithm of *Buscombe et al.* [2010] for mean particle size has been proposed for granular sections with known (or measurable) void fraction. Such a measure has potential applicability in thin section, microscopic or holographic imaging of particles. These methods have no tunable parameters or empirically derived coefficients, so they should be broadly universal in appropriate application. However, empirical corrections increase accuracy.

[57] Images of the surfaces of both real and simulated (using the methods in part 1) granular material have been used to evaluate the new methods for use in the absence of an apparent void fraction. Using real samples, ensemble RMS errors are 30.9% and 27.4% for standard deviation of particle sizes without and with bias-correction, respectively. The bias-correction accounts for sediment population- or camera/lighting-specific factors and takes the simplest possible form, which is a slope adjustment of estimates. Smaller RMS errors would be expected if both a slope and intercept correction were used. Using the simulated surfaces with no apparent void fraction, ensemble RMS errors are 14.9% and 13.9% for mean and standard deviation of particle sizes, respectively, without bias-correction.

[58] Images of 2D sections through simulated granular volumes have been used to evaluate the methods for estimating mean and standard deviation of particle sizes when the volume has a known apparent (areal) void fraction. Ensemble RMS error for sorting is 10.4% without bias-correction. The same for mean size is a rather unsatisfactory 36%, but reduces to 9% if a correction is applied. This empirical correction was found using void fraction alone, suggesting that the method for mean particle size is incorporating void fraction correctly only to first-order, and that an extra term is needed when void fraction is less than approximately 0.5.

[59] Simulations have also been used to quantify the approximate discrepancy between estimates of mean particle size from the visible layers of the surface and that of the volume as a whole. These discrepancies arise because the surface has a porosity which is not apparent because of particle packing which makes them appear to overlap. It was found that particles, either whole or in part, are visible from an overhead perspective to a depth of approximately 4 mean particle diameters. Estimates of mean particle size from automated methods such as here (also *Rubin* [2004] and *Buscombe et al.* [2010]) are likely to be smaller than the population mean due to the effects of perspective, whereby parts of overlapped particles in the image are measured as well as whole particles.

[60] Finally, a simple analytical formula has been derived for the one-step pore-particle transition probability matrix, estimated from the easily computed autocorrelogram, from which void fraction on a granular material can be estimated directly. This model gives highly accurate predictions of bulk void fraction yet imperfect predictions of pore-particle transitions.

## Appendix A: Correcting Estimates of Mean Particle Size in Granular Material With Apparent Porosity

[61] In order to find the functional form of the error in estimates of mean particle size in granular material (with an apparent porosity, $\phi$) with respect to $\phi$, we use linear optimization to solve the problem:

$$\min_x \frac{1}{2}(Cx - d)^2. \tag{A1}$$

The curve is forced to pass through points $x_0$ and $y_0$ by subjecting $x$ to an inequality constraint $x_{eq}x = y_{eq}$, where $x_{eq}$ is the Vandermonde matrix for $x_0$, and $y_{eq} = y_0$. Matrix $C$ is the Vandermonde matrix for $\phi$; and $d$ are the target values. When $d = (O - E)/O$, and $y_0 = 0$ and $x_0 = \pi$. Here, when $d = (O - E)$, $x_{eq} = 0$ and $y_{eq} = -180$ ($-\pi$, estimated visually as the intercept). We have used least squares order 3 (cubic polynomial).

[62] **Acknowledgments.** Thanks to Jon Warrick for early discussions regarding sorting coefficient estimation, and for the Elwha image data used to verify the new technique. Thanks to Hank Chezar for designing and constructing the bed sediment cameras used to take images for this study, and Parker Allwardt for carrying out numerous manual point counts on some of those images, from which we derived population statistics. This work was started while the first author was a postdoctoral research fellow at the United States Geological Survey (Western Coastal and Marine Geology) and the Institute of Marine Studies, University of California, Santa Cruz. Thanks to Alex Nimmo-Smith, Chris Sherwood, Jingping Xu, Patrice Carbonneau, and two anonymous reviewers for their constructive comments. A program (My Automated Grain size from Images Code - 'MAGIC') is available from http://walrus.wr.usgs.gov/seds/grainsize/code.html.

## References

Barnard, P., D. Rubin, J. Harney, and N. Mustain (2007), Field test comparison of an autocorrelation technique for determining grain size using a digital 'beachball' camera versus traditional methods, *Sediment. Geol.*, *201*, 180–195.

Bartlett, M. (1946), On the theoretical specification and sampling properties of autocorrelated time series, *J. R. Stat. Soc.*, *8*, 27–41.

Berne, B. J., and R. Pecora (2000), *Dynamic Light Scattering*, Dover, New York.

Buscombe, D. (2008), Estimation of grain-size distributions and associated parameters from digital images of sediment, *Sediment. Geol.*, *210*, 1–10.

Buscombe, D., and G. Masselink (2009), Grain size information from the statistical properties of digital images of sediment, *Sedimentology*, *56*, 421–438.

Buscombe, D., and D. M. Rubin (2012), Advances in the simulation and automated measurement of well-sorted granular material: 1. Simulation, *J. Geophys. Res.*, *117*, F02002, doi:10.1029/2011JF001974.

Buscombe, D., D. M. Rubin, and J. A. Warrick (2010), A universal approximation to grain size from images of non-cohesive sediment, *J. Geophys. Res.*, *115*, F02015, doi:10.1029/2009JF001477.

Cressie, N. (1993), *Statistics for Spatial Data*, Wiley, New York.

Cresswell, A. W., and M. E. Barton (2003), Direct shear tests on an uncemented, and a very slightly cemented, locked sand, *Q. J. Eng. Geol. Hydrogeol.*, *36*, 119–132.

Davis, J. C. (1986), *Statistics and Data Analysis in Geology*, 2nd ed., Wiley, New York.

DeGroot, D., and G. Baecher (1993), Estimating autocovariances of in-situ soil properties, *J. Geotech. Eng.*, *119*, 147–166.

Du, Q., V. Faber, and M. Gunzburger (1999), Centroidal Voronoi Tessellations: Applications and algorithms, *SIAM Rev.*, *41*, 637–676.

Fenton, G. (1999), Random field modeling of CPT data, *J. Geotech. Geoenviron. Eng.*, *125*, 486–498.

Fieller, N., E. Flenley, and W. Olbricht (1992), Statistics of particle size data, *Appl. Stat.*, *41*, 127–146.

Goff, J. A., and T. H. Jordan (1988), Stochastic modeling of seafloor morphology: Inversion of sea beam data for second-order statistics, *J. Geophys. Res.*, *93*, 13,589–13,608, doi:10.1029/JB093iB11p13589.

Gorecki, C. (1989), Optical sizing by Fourier transformation, *J. Opt.*, *20*, 25–29.

Grady, D. (1990), Particle size statistics in dynamic fragmentation, *J. Appl. Phys.*, *68*, 6099–6105.

Graham, G., and W. Nimmo-Smith (2010), The application of holography to the analysis of size and settling velocity of suspended cohesive sediments, *Limnol. Oceanogr. Methods*, *8*, 1–15.

Halton, J. (1960), On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals, *Numer. Math.*, *2*, 84–90.

Holliger, K., A. R. Levander, and J. A. Goff (1993), Stochastic modeling of the reflective lower crust: Petrophysical and geological evidence from the Ivrea zone (Northern Italy), *J. Geophys. Res.*, *98*, 11,967–11,980, doi:10.1029/93JB00351.

Kemeny, J., and J. Snell (1960), *Finite Markov Chains*, Van Nostrand, Princeton, N. J.

Koutsourelakis, P.-S., and G. Deodatis (2005), Simulation of binary random fields with applications to two-phase random media, *J. Eng. Mech.*, *131*, 397–412.

Liang, Z., C. Fernandes, F. Magnani, and P. Philippi (1998), A reconstruction technique for three-dimensional porous media using image analysis and Fourier transforms, *J. Petrol. Sci. Eng.*, *21*, 273–283.

Momota, M., H. Miike, and H. Hashimoto (1994), Measuring particle size distribution by digital image processing with inverse Fourier-Bessel transformation, *Jpn. J. Appl. Phys.*, *33*, 1189–1194.

Muche, L., and D. Stoyan (1992), Contact and chord distributions of the Poisson-Voronoi tessellation, *J. Appl. Probab.*, *29*, 467–471.

Nikora, V., D. Goring, and B. Biggs (1998), On gravel-bed roughness characterization, *Water Resour. Res.*, *34*, 517–527, doi:10.1029/97WR02886.

Oppenheim, A., R. Schafer, and J. Buck (1999), *Discrete-Time Signal Processing*, 2nd ed., Prentice-Hall, New York.

Pereira, F., and M. Gharib (2004), A method for three-dimensional particle sizing in two-phase flows, *Meas. Sci. Technol.*, *15*, 2029–2038.

Preston, F. W., and J. C. Davis (1976), Sedimentary porous materials as a realization of a stochastic process, in *Random Processes in Geology*, edited by D. F. Merriam, pp. 63–86, Springer, New York.

Priestley, M. (1981), *Spectral Analysis and Time Series*, Academic, London.

Prince, C., R. Ehrlich, and Y. Anguy (1995), Analysis of spatial order in sandstones II: Grain clusters, packing, flaws, and the small-scale structure of sandstones, *J. Sediment. Res.*, *65*, 13–28.

Ripley, B. (1981), *Spatial Statistics*, Wiley, London.

Rubin, D. (2004), A simple autocorrelation algorithm for determining grain size from digital images of sediment, *J. Sediment. Res.*, *74*, 160–165.

Rubin, D., H. Chezar, J. Harney, D. Topping, T. Melis, and C. Sherwood (2007), Underwater microscope for measuring spatial and temporal changes in bed-sediment grain size, *Sediment. Geol.*, *202*, 402–408.

Sen, Z. (1978), Autorun analysis of hydrologic time series, *J. Hydrol. 36*, 75–85.

Sen, Z. (1984), Autorun analysis of sedimentary porous materials, *Math. Geol.*, *16*, 449–463.

Taylor, G. (1938), The spectrum of turbulence, *Proc. R. Soc. London, ser. A*, *164*, 476–490.

Torabi, A., H. Fossen, and B. Alaei (2008), Application of spatial correlation functions in permeability estimation of deformation bands in porous rocks, *J. Geophys. Res.*, *113*, B08208, doi:10.1029/2007JB005455.

Uzielli, M., G. Vannucchi, and K. Phoon (2005), Random field characterisation of stress-normalised cone penetration testing parameters, *Geotechnique*, *55*, 3–20.

Warrick, J. A., D. M. Rubin, P. Ruggiero, J. Harney, A. E. Draut, and D. Buscombe (2009), Cobble cam: Grain-size measurements of sand to boulder from digital photographs and autocorrelation analyses, *Earth Surf. Processes Landforms*, *34*, 1811–1821.

D. Buscombe, School of Marine Science and Engineering, University of Plymouth, Plymouth PL4 8AA, UK. (daniel.buscombe@plymouth.ac.uk)

D. M. Rubin, United States Geological Survey, Santa Cruz, CA 95060, USA. (drubin@usgs.gov)